

Improved bounds on number fields of small degree

Theresa C. Anderson* Ayla Gafni† Kevin Hughes‡
Robert J. Lemke Oliver§ David Lowry-Duda¶ Frank Thorne||
Jiuya Wang** Ruixiang Zhang††

Received 9 September 2022; Revised 26 October 2023; Published 23 December 2024

Abstract: We study the number of degree n number fields with discriminant bounded by X . In this article, we improve an upper bound due to Schmidt on the number of such fields that was previously the best known upper bound for $6 \leq n \leq 94$.

1 Introduction and outline of paper

Let $N_n(X) := \#\{K/\mathbb{Q} : [K : \mathbb{Q}] = n, |\text{Disc}(K)| \leq X\}$ count the number of degree n number fields over \mathbb{Q} with bounded discriminant of size X . We consider all extensions as being inside a fixed algebraic closure $\overline{\mathbb{Q}}$ of \mathbb{Q} . A classical theorem of Hermite shows that $N_n(X)$ is finite, and an optimized and modern version of Hermite's argument due to Schmidt [Sch95] shows that $N_n(X)$ may be bounded by $O_n(X^{\frac{n+2}{4}})$. It is

*Supported by NSF DMS-2231990, DMS-1954407, and CAREER DMS-2237937

†Supported by Simons Collaboration Grant No. 946710

‡Supported by the Additional Funding Programme for Mathematical Sciences, delivered by EPSRC (EP/V521917/1) and the Heilbronn Institute for Mathematical Research

§Supported by NSF DMS-2200760 and a Simons Fellowship in Mathematics

¶Supported by the Simons Collaboration in Arithmetic Geometry, Number Theory, and Computation via the Simons Foundation Grant No. 546235

||Supported by NSF DMS-2101874 and a Simons Collaboration Grant No. 586594

**Supported by NSF DMS-2201346

††Supported by NSF DMS-1856541 and NSF DMS-2207281

conjectured that the actual sharp upper bound is $O_n(X)$; thus improvements to the Schmidt bound, and the classical work of Hermite, have attracted a lot of attention.

For $n \leq 5$, precise asymptotic formulas for $N_n(X)$ are known [DH69, CDyDO02, Bha05, Bha10]. For large n , there are substantial improvements over the Schmidt bound, beginning with work of Ellenberg and Venkatesh [EV06], which was subsequently improved in work of Couveignes [Cou20] and Lemke Oliver and Thorne [LOT22]. For $n \geq 95$, the best bound is that of [LOT22], namely $N_n(X) \ll_n X^{c(\log n)^2}$, where $c \leq 1.564$ is an explicitly computable constant. While this improvement is substantial as $n \rightarrow \infty$, in the complementary regime $6 \leq n \leq 94$, the Schmidt bound has remained the best available for 25 years.

Our main theorem gives a general improvement over the Schmidt bound.

Theorem 1.1. *For each $n \geq 6$ and every $\epsilon > 0$, we have $N_n(X) \ll_{n,\epsilon} X^{\frac{n+2}{4} - \frac{1}{4n-4} + \epsilon}$.*

We do not expect that the saving $\frac{1}{4n-4}$ is optimal for our method, but to extract a further saving from the work below would take significant care. In fact, our work is inspired by work of Bhargava, Shankar, and Wang [BSW22b] in a different context, and simultaneously with our work, an improvement of (essentially) $\frac{1}{2n-2}$ over Schmidt’s bound was independently obtained by the same authors in [BSW22a]. Their independent work bears some similarity with our approach, but remarkably the two approaches differ greatly in multiple ways. Therefore, we expedite some pieces of our approach, while focusing on other pieces where our techniques are different and which we expect to have other applications in different settings.

In §2, we summarize Schmidt’s approach of estimating $N_n(X)$ by counting monic polynomials $f(x) = x^n + c_1x^{n-1} + \cdots + c_n$ where each c_i is bounded by $O(X^{\frac{i}{2n-2}})$. We slightly restructure Schmidt’s method in order to not restrict to polynomials of trace 0. To improve upon Schmidt’s bound, we separately count polynomials with “particularly small” discriminants and polynomials with “particularly squarefull” discriminants in comparison to the discriminants of the fields they cut out. Our method is easier to describe in the context of Schmidt’s method; thus, we defer further description until §2.1. Roughly speaking, we study polynomials with “particularly small” discriminant in §4 and polynomials with “particularly squarefull” discriminant in §6 and §7. Inspired by [ST23], we bound both the archimedean and nonarchimedean local density of polynomials of small discriminant in Proposition 4.1; these densities are important in §6 and §7. These densities bear some resemblance to the ‘Main Heuristic assumption’ in [ST23], but Proposition 4.1, though natural, appears to be novel. Due to the significant blend of analytic techniques developed to work with polynomials in different contexts, we devote an early section, Section 3, to their exposition.

It is in §5 and §6 that our approach differs most substantially from [BSW22b] and [BSW22a]. In §5, which we regard as our most novel, we study the nonarchimedean Fourier transform of the set of polynomials whose discriminant is divisible by a large power of a prime. The main results of this section are somewhat technical statements about the support and size of this Fourier transform, but their proofs rely on a beautiful and substantial structure of the discriminant that is revealed upon taking its Fourier transform. We expect that such Fourier transforms, and the underlying structure of the discriminant so revealed, may play a role in other problems beyond that considered in Theorem 1.1, for example perhaps in generalizing the ideas from [ST23]. We use the considerations of §5 to show in §6 bounds for polynomials whose discriminants are “powerful”: that is, they admit prime factors to large powers. Section 3 describes the various techniques, such as change of variables, that permeate the next three sections. We hope this

reference section facilitates the reading of the rest of the paper. Finally, in §7, we study polynomials whose discriminants have a large factor that is not powerful, incorporating the framework of [BSW22b]. The specific prime factorization of the large factor dividing the polynomial discriminant determines which polynomials are treated using §6 or §7, and is described in the next section.

2 Outline of approach

Let $\mathcal{F}_n(X)$ denote the set of fields K/\mathbb{Q} with degree n and discriminant satisfying $|\text{Disc}(K)| \leq X$. Note that $N_n(X) = \#\mathcal{F}_n(X)$, our desired count. The Schmidt bound [Sch95] asserts that $\#\mathcal{F}_n(X) \ll_n X^{\frac{n+2}{4}}$. The key to Schmidt's approach is the following.

Lemma 2.1. *Let $K \in \mathcal{F}_n(X)$ and assume K is primitive, i.e. that it has no proper nontrivial subfields. Then there is a monic polynomial $f(x) \in \mathbb{Z}[x]$ with trace 0, i.e. $f(x) = x^n + c_2x^{n-2} + \dots + c_n$, for which $K \simeq \mathbb{Q}[x]/\langle f(x) \rangle$ and where each c_i satisfies $|c_i| \ll_n X^{\frac{i}{2n-2}}$.*

Proof. In the Minkowski embedding, the trace 0 elements of \mathcal{O}_K form a rank $n-1$ lattice with covolume $\asymp_n |\text{Disc}(K)|^{1/2}$. The $n-1$ Minkowski minima for this lattice satisfy $\lambda_1 \cdots \lambda_{n-1} \asymp |\text{Disc}(K)|^{1/2}$, so $\lambda_1 \ll_n |\text{Disc}(K)|^{1/2(n-1)}$. In particular, there is a non-zero element $\alpha \in \mathcal{O}_K$ with trace 0 whose height is $O_n(|\text{Disc}(K)|^{\frac{1}{2n-2}})$. Since α has trace 0 and $\alpha \neq 0$, it follows that $\alpha \notin \mathbb{Q}$. By our assumption that K admits no interesting subfields, it follows that $\mathbb{Q}(\alpha) = K$. The minimal polynomial of α then satisfies the conclusion of the lemma. \square

There are $O_n(X^{\frac{n+2}{4}})$ polynomials of the type produced in Lemma 2.1. It thus follows that the number of primitive fields contained in $\mathcal{F}_n(X)$ is also $O_n(X^{\frac{n+2}{4}})$.

Remark 2.2. To get a bound on non-primitive fields too, Schmidt (essentially) uses this argument, but over a moving subfield. His bound actually improves for the number of imprimitive fields: if d_{\max} is the largest proper divisor of n , his bound on imprimitive fields in $\mathcal{F}_n(X)$ is $X^{\frac{d_{\max}+2}{4}}$. This is smaller than the bound claimed in Theorem 1.1, so it suffices to prove the claimed bound holds for the set $\mathcal{F}_n^{\text{prim}}(X)$ of primitive field extensions.

It is possible to rework Schmidt's argument slightly to make it marginally more convenient to invoke the results of the earlier work of Bhargava, Shankar, and Wang [BSW22b]. First, we consider polynomials without trace 0.

Define the *height* H of a polynomial $f(x) = x^n + c_1x^{n-1} + \dots + c_n$ by $H = \max |c_i|^{1/i}$. We will frequently abuse notation by using the phrase ‘polynomials of height H ’ to mean polynomials with height bounded by H .

Lemma 2.3. *Let $n \geq 2$. There is a constant $C_n > 0$ such that each field $K \in \mathcal{F}_n^{\text{prim}}(X)$ is cut out by $\gg_n X^{\frac{1}{2n-2}}$ polynomials $f(x) \in \mathbb{Z}[x]$ of the form $f(x) = x^n + c_1x^{n-1} + \dots + c_n$, where each c_i satisfies $|c_i| \leq C_n^i X^{\frac{i}{2n-2}}$.*

Proof. By the proof of Lemma 2.1, each $K \in \mathcal{F}_n^{\text{prim}}(X)$ has an element $\alpha \in \mathcal{O}_K$ with trace 0 and height bounded by $H_0 := B_n X^{\frac{1}{2n-2}}$, for a positive constant B_n . Then there are at least H_0 elements in \mathcal{O}_K of height at most $2H_0$, corresponding to the translates $\alpha + k$ for $|k| \leq H_0$, $k \in \mathbb{Z}$. The minimal polynomial $m(x) = x^n + c_1x^{n-1} + \dots + c_n$ of such an element satisfies $|c_i| \ll_n H_0^i$, and the result follows. \square

For an admissible choice of C_n in Lemma 2.3, set $H := C_n X^{\frac{1}{2n-2}}$. There are $O_n(H^{\frac{n^2+n}{2}})$ polynomials of height H , with each primitive field cut out by at least $\gg_n H$ different polynomials. Dividing by this minimum multiplicity, we find that there must be no more than $O_n(H^{\frac{n^2+n-2}{2}}) = O_n(X^{\frac{n+2}{4}})$ different fields produced, again recovering Schmidt's bound.

Remark 2.4. The argument of Lemma 2.3 is a bit ad hoc, but its conclusion is likely close to optimal, apart from constants depending on n . For “typical” fields $K \in \mathcal{F}_n(X)$, we should expect the multiplicity with which K is cut out to be $\asymp_n H^n / \sqrt{|\text{Disc}(K)|}$. If $H = X^{\frac{1}{2n-2}} \approx |\text{Disc}(K)|^{\frac{1}{2n-2}}$, this expected multiplicity is about H . The proof of Lemma 2.3 thus doesn't lose much, and its proof will be convenient in the next section.

2.1 Strategy and Proof of Main Theorem

While Lemmas 2.1 and 2.3 guarantee that each field $K \in \mathcal{F}_n^{\text{prim}}(X)$ is cut out by a polynomial of height $H \ll_n X^{\frac{1}{2n-2}}$, it is not the case that every such polynomial cuts out a field of discriminant at most X . In particular, for “typical” $f(x)$ of height H , we have $\text{disc}(f) \approx H^{n^2-n} = X^{\frac{n}{2}}$, and this is typically also the order of the discriminant of the field cut out by f . We should thus expect the relevant polynomials attached to $K \in \mathcal{F}_n^{\text{prim}}(X)$ to be exceptional in one of two ways: either the discriminant of f is unusually small, or the discriminant of the field cut out by f is much smaller than the discriminant of f . In the latter case, the ratio of the two discriminants is the square of the index $[\mathcal{O}_K : \mathbb{Z}[\alpha]]$ where α is a root of f ; we call this integer the index of the polynomial f , denoted $\text{index}(f)$.

The condition that the discriminant is small is global, while the condition that the index is large is local. However, since both conditions are preserved under translation, it follows from the proof of Lemma 2.3 that each $K \in \mathcal{F}_n^{\text{prim}}(X)$ will still be cut out by $\gg_n H$ of these exceptional polynomials. Thus, our main task is to bound the number of exceptional polynomials.

Proof of Theorem 1.1

We now describe the proof of our main theorem more concretely.

By Lemma 2.3, to bound $K \in \mathcal{F}_n^{\text{prim}}(X)$ it suffices to bound the number of polynomials f with height up to $H \asymp_n X^{\frac{1}{2n-2}}$ that cut out a field with discriminant $X \asymp_n H^{2n-2}$, where each field is counted with multiplicity of order at least H .

In Section 4, we prove the following corollary, bounding the the number of polynomials with small discriminant.

Corollary 2.5. *Let $n \geq 3$ and let H be sufficiently large in terms of n . The number of polynomials $f(x) \in \mathbb{Z}[x]$ of the form $f(x) = x^n + c_1 x^{n-1} + \cdots + c_n$ with $|c_i| \leq H^i$ and $|\text{disc}(f)| \leq H^{n^2-n-2}$ is $O_n(H^{\frac{n^2+n}{2}-1})$.*

If $f(x)$ is irreducible and cuts out the field K , then the discriminants of f and K are related through the equation $\text{disc}(f) = \text{Disc}(K)[\mathcal{O}_K : \mathbb{Z}[\alpha]]^2$, where α is a root of f over K . As indicated above, we refer to the index $[\mathcal{O}_K : \mathbb{Z}[\alpha]]$ as the index of f , which we denote by $\text{index}(f)$. Corollary 2.5 implies that with at most $O_n(H^{\frac{n^2+n}{2}-1})$ exceptions,

$$\text{index}(f)^2 \cdot \text{Disc}(K) = \text{disc}(f) > H^{n^2-n-2}.$$

As $\text{Disc}(K) \leq X \asymp_n H^{2(n-1)}$, each of these polynomials has large index bounded from below by

$$\text{index}(f) \gg_n H^{\frac{n^2-n-2}{2}-(n-1)} = H^{\frac{n(n-3)}{2}}. \quad (2.1)$$

Thus each of these polynomials has discriminant divisible by a large square. To bound the number of such polynomials, we consider two subcases: when the radical of $\text{index}(f)$ is small or large. (Recall that the radical of an integer is the product of its prime divisors.) In Section 6, we prove the following bound on the number of polynomials with large index, but small index radical.

Theorem 2.6. *Let $n \geq 6$ and fix any $\epsilon > 0$. Then for any $H \geq 1$, the number of polynomials $f(x) \in \mathbb{Z}[x]$ of degree n and height H for which $\text{rad}(\text{index}(f)) < H^{1-\epsilon}$ but $\text{index}(f) > H^{\frac{n(n-3)}{2}}$ is $O_{n,\epsilon}(H^{\frac{n^2+n}{2}-\frac{4}{3}-\frac{4}{n}+\epsilon} + H^{\frac{n^2+n}{2}-\frac{2n}{3}+3+\epsilon})$.*

Note that for all $n \geq 8$, the first term above is larger, and therefore governs our overall bound. However, for $n = 6, 7$ the second term dominates.

In Section 7, we bound the number of polynomials with large index radical.

Theorem 2.7. *For any $n \geq 3$, any $H \geq 1$, any $\epsilon > 0$, and any $M \geq 1$, the number of polynomials f of degree n , height H , and where $m^2 \mid \text{disc } f$ for some squarefree $m \geq M$ is*

$$O_{n,\epsilon} \left(\frac{H^{\frac{n^2+n}{2}}}{\sqrt{M}} + H^{\frac{n^2+n}{2}-\frac{1}{2}+\epsilon} \right).$$

Theorem 2.6 implies that the number of polynomials of height H , index bounded below by (2.1), and with $\text{rad}(\text{index}(f)) < H^{1-\epsilon}$ is $O_{n,\epsilon}(H^{\frac{n^2+n}{2}-1+\epsilon})$. All remaining uncounted polynomials have $\text{rad}(\text{index}(f)) \geq H^{1-\epsilon}$, and thus have a squarefree divisor m of size at least $H^{1-\epsilon}$ such that $m^2 \mid \text{disc}(f)$. Taking $M = H^{1-\epsilon}$ in Theorem 2.7 shows that there are at most $H^{\frac{n^2+n}{2}-\frac{1}{2}+\epsilon}$ such polynomials. Combining these bounds, we find that

$$H \cdot \#\mathcal{F}_n^{\text{prim}}(X) \ll_{n,\epsilon} H^{\frac{n^2+n}{2}-\frac{1}{2}+\epsilon},$$

which completes the proof of Theorem 1.1 using $H \asymp_n X^{\frac{1}{2n-2}}$. \square

3 Local fields, étale algebras and mass formulas

Changes of variables between étale algebras, polynomial roots and polynomial coefficients will be used throughout this paper, so we introduce this section as a useful reference. As such, the terminology here is self-contained.

Let v be a place of \mathbb{Q} . Throughout most of this paper, we will be concerned with the properties of the set of monic, degree n polynomials $f \in \mathbb{Z}[x]$ for which $|\text{disc}(f)|_v$ is small in some suitable sense. We approach these questions locally, viewing our polynomials (perhaps after a suitable change of variables) as lying inside $\mathcal{O}_{\mathbb{Q}_v}[x]$, where $\mathcal{O}_{\mathbb{Q}_v} = \mathbb{Z}_v$ if v is finite and $\mathcal{O}_{\mathbb{Q}_v} = [-1, 1]$ if $v = \infty$ (note that this selection of $[-1, 1]$ instead of \mathbb{R} may be nonstandard, but is useful in our context). Rather than study the discriminant of f as a polynomial in its coefficients $\mathbf{c} \in \mathcal{O}_{\mathbb{Q}_v}^n$, we prefer to study the discriminant by means of its simpler

expression in terms of the roots of f . These roots may be thought of as an n -tuple $\lambda \in \overline{\mathbb{Q}_v}^n$ for a fixed choice of algebraic closure $\overline{\mathbb{Q}_v}$, but the locus of points inside $\overline{\mathbb{Q}_v}^n$ actually arising as roots of polynomials in $\mathcal{O}_{\mathbb{Q}_v}[x]$ appears difficult to access directly. Thus, we pass through an intermediate space, the étale algebra associated with f .

Let $\mathcal{P}_n(\mathcal{O}_{\mathbb{Q}_v})$ be the space of monic, degree n polynomials with coefficients in $\mathcal{O}_{\mathbb{Q}_v}$. Given a separable polynomial $f \in \mathcal{P}_n(\mathcal{O}_{\mathbb{Q}_v})$, the quotient $\mathbb{Q}_v[x]/f$ is a degree n étale algebra K_v over \mathbb{Q}_v . This étale algebra is equipped with a natural element (namely, the image of x), and, conversely, given an element $\alpha \in K_v$, the polynomial $\text{Nm}(x - \alpha)$ will be associated with α under this map. Now, for any v , there are only finitely many degree n étale algebras K_v over \mathbb{Q}_v . Thus, the image of the map $f \mapsto x \in \mathbb{Q}_v[x]/f$ naturally decomposes as a disjoint union over the different étale algebras. Moreover, if v is finite, then the image of f in fact lands in \mathcal{O}_{K_v} , the ring of integers of K_v . If $v = \infty$, then we define \mathcal{O}_{K_v} to be the compact closure of the image.

Given $\alpha \in \mathcal{O}_{K_v}$, the roots of the associated polynomial $f_\alpha = \text{Nm}(x - \alpha)$ are easily determined. Explicitly, if we write $K_v = \bigoplus_{i=1}^r F_i$ with each F_i a (field) extension of \mathbb{Q}_v , then any $\alpha \in \mathcal{O}_{K_v}$ may be expressed as $\alpha = (\alpha_1, \dots, \alpha_r)$ with each $\alpha_i \in F_i$. The roots of f_α are then the images of each α_i under the $[F_i : \mathbb{Q}_v]$ embeddings of F_i into our fixed choice $\overline{\mathbb{Q}_v}$. Thus, while we are in principle interested in performing computations in the space of roots $\lambda \in \overline{\mathbb{Q}_v}^n$, in practice, we instead perform these computations inside étale algebras.

In carrying out the computations to follow, we find it convenient to fix a coordinatization of \mathcal{O}_{K_v} when v is finite. In particular, if $K_v = \bigoplus_{i=1}^r F_i$ as above, then by choosing an integral basis $\omega_{i,1}, \dots, \omega_{i,[F_i:\mathbb{Q}_v]}$ for each F_i , the collection $\omega_{1,1}, \dots, \omega_{r,[F_r:\mathbb{Q}_v]}$ is an integral basis for K_v . This identifies \mathcal{O}_{K_v} with $\mathcal{O}_{\mathbb{Q}_v}^n$, whose generic element we denote \mathbf{a} . The transformation from the coordinates \mathbf{a} to the roots λ is linear, with block diagonal matrix $M_{K_v} = \bigoplus_{i=1}^r M_{F_i}$, where

$$M_{F_i} = \begin{pmatrix} \iota_1(\omega_{i,1}) & \dots & \iota_1(\omega_{i,[F_i:\mathbb{Q}_v]}) \\ \vdots & & \vdots \\ \iota_{[F_i:\mathbb{Q}_v]}(\omega_{i,1}) & \dots & \iota_{[F_i:\mathbb{Q}_v]}(\omega_{i,[F_i:\mathbb{Q}_v]}) \end{pmatrix},$$

with $\iota_1, \dots, \iota_{[F_i:\mathbb{Q}_v]}$ denoting the embeddings of F_i into $\overline{\mathbb{Q}_v}$. Notice that $|\det M_{F_i}|_v = |\text{disc}(F_i)|_v^{1/2}$ by definition of the discriminant, hence $|\det M_{K_v}|_v = |\text{disc}(K_v)|_v^{1/2}$.

Closing the loop, the roots λ of a polynomial $f \in \mathcal{P}_n(\mathcal{O}_{\mathbb{Q}_v})$ naturally determine its coefficients \mathbf{c} by means of the elementary symmetric polynomials. Namely, given the roots $\lambda = (\lambda_1, \dots, \lambda_n) \in \overline{\mathbb{Q}_v}^n$ of a monic polynomial f , define the change of variable λ to $\sigma = (\sigma_1, \dots, \sigma_n) \in \overline{\mathbb{Q}_v}^n$ via

$$\sigma_i(\lambda) = \sum_{\substack{S \subset \{1, \dots, n\} \\ |S|=i}} \prod_{j \in S} \lambda_j$$

for $i = 1, \dots, n$. This change of variable yields the polynomial $f(x) = x^n - \sigma_1 x^{n-1} + \dots + (-1)^k \sigma_k x^{n-k} + \dots + (-1)^n \sigma_n$. Each σ_i is the i -th elementary symmetric polynomial in λ_i .

Figure 1 summarizes the different spaces, maps between them, and variable naming conventions we adopt.

At this stage, it only remains to discuss the change of measure between the various spaces, in particular between polynomials and étale algebras. First, the space $\mathcal{P}_n(\mathcal{O}_{\mathbb{Q}_v})$ is equipped with a natural

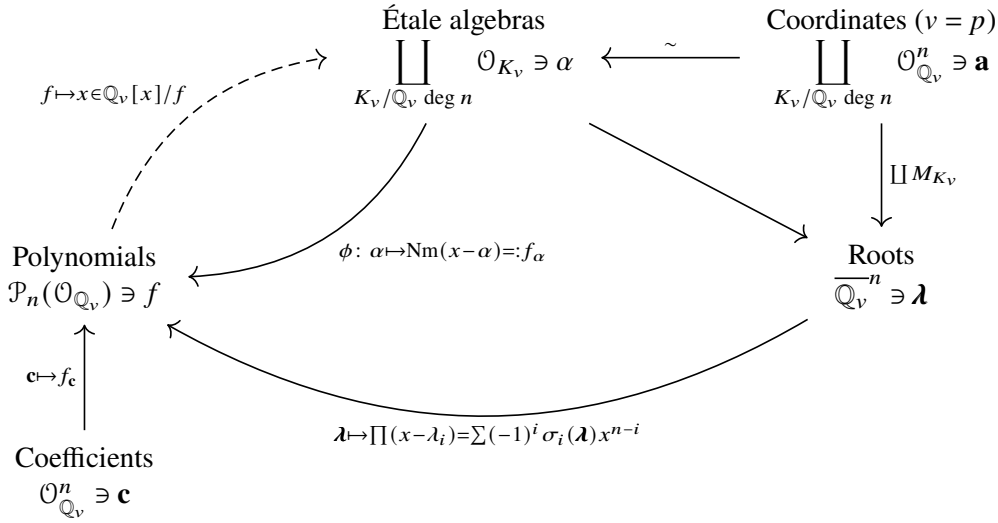


Figure 1: Maps between spaces of polynomials, étale algebras, and roots. The dashed arrow indicates this map is defined for separable polynomials.

Haar/Lebesgue measure ν by means of the coefficient isomorphism $\mathcal{O}_{\mathbb{Q}_v}^n \rightarrow \mathcal{P}_n(\mathcal{O}_{\mathbb{Q}_v})$; if ν is finite, we normalize the measure ν so that $\nu(\mathcal{P}_n(\mathcal{O}_{\mathbb{Q}_v})) = 1$. Each étale algebra K_v is also equipped with a natural Haar/Lebesgue measure that we denote μ , normalized if ν is finite so that $\mu(\mathcal{O}_{K_v}) = 1$. The fundamental result we use is due to Serre [Ser08, Lemma 3] and Shankar–Tsimmerman [ST23, Lemma 2.2].

Lemma 3.1. *With notation as above, we have $\phi^* \nu = |\text{disc}(K_v)|_v^{1/2} |\text{disc}(f_\alpha)|_v^{1/2} \mu$. In particular, for any ν -integrable function ψ on $\mathcal{P}_n(\mathcal{O}_{\mathbb{Q}_v})$, we have*

$$\int_{\mathcal{P}_n(\mathcal{O}_{\mathbb{Q}_v})} \psi(f) d\nu = \sum_{K_v/\mathbb{Q}_v \text{ deg. } n} \frac{|\text{disc}(K_v)|_v^{1/2}}{|\text{Aut}(K_v)|} \int_{\mathcal{O}_{K_v}} |\text{disc}(f_\alpha)|_v^{1/2} \psi(f_\alpha) d\mu.$$

4 The density of polynomials with small discriminant

We begin by proving the following proposition on the density of polynomials over some completion \mathbb{Q}_v of \mathbb{Q} with small discriminant. This result in the case that $\mathbb{Q}_v = \mathbb{R}$ will be the key ingredient in the proof of Corollary 2.5, and the ideas behind the proof in the case that \mathbb{Q}_v is non-archimedean will be the key ingredient in the proof of Theorem 2.6.

Proposition 4.1. *Let $n \geq 2$ and let \mathbb{Q}_v be a completion of \mathbb{Q} . Let $|\cdot|_v$ be the associated absolute value and let μ_v be the associated Haar measure on \mathbb{Q}_v , normalized in the case that ν is finite so that the total measure of \mathbb{Z}_v is 1 and so that μ_v agrees with usual Lebesgue measure in the case that ν is infinite. For $\mathbf{c} \in \mathbb{Q}_v^n$, let $f_{\mathbf{c}}(x) := x^n + c_1 x^{n-1} + \cdots + c_n$.*

Then for any $\delta \in (0, 1)$, there holds

$$\nu(\{\mathbf{c} \in \mathbb{Q}_v^n : |c_i|_v \leq 1 \text{ for all } i \text{ and } |\text{disc}(f_{\mathbf{c}}(x))|_v \leq \delta\}) \ll_n \delta^{\frac{1}{2} + \frac{1}{n}}, \quad (4.1)$$

where ν denotes the product measure on \mathbb{Q}_v^n .

Proof. Let $\mathbf{1}_\delta(f)$ be the characteristic function for the set of polynomials where $|\text{Disc}(f)|_v \leq \delta$. Then applying the change of variables discussed in Section 3, we find

$$\begin{aligned} & \nu(\{\mathbf{c} \in \mathbb{Q}_v^n : |c_i|_v \leq 1 \text{ for all } i \text{ and } |\text{disc}(f_{\mathbf{c}}(x))|_v \leq \delta\}) \\ &= \sum_{[K_v:\mathbb{Q}_v]=n} \frac{|\text{Disc}(K_v)|_v^{1/2}}{|\text{Aut}(K_v)|} \int_{\mathcal{O}_{K_v}} |\text{Disc}(f_\alpha)|_v^{1/2} \mathbf{1}_\delta(f_\alpha) d\mu(\alpha) \\ &\leq \delta^{1/2} \sum_{[K_v:\mathbb{Q}_v]=n} \frac{|\text{Disc}(K_v)|_v^{1/2}}{|\text{Aut}(K_v)|} \int_{\mathcal{O}_{K_v}} \mathbf{1}_\delta(f_\alpha) d\mu(\alpha), \end{aligned}$$

where f_α denotes the characteristic polynomial of α , and where, in the inequality in the second line, we used the bound $|\text{Disc}(f_\alpha)|_v = |\text{Disc}(\alpha)|_v \leq \delta$, valid for any α for which $\mathbf{1}_\delta(f_\alpha) \neq 0$. Note in particular that any polynomial with a repeated root has discriminant 0, but that the set of such polynomials has measure 0 and thus does not affect the integral.

We now estimate the integral, beginning with the case where $K_v \simeq \mathbb{Q}_v^n$ is the totally split algebra. In this case, $|\text{Disc}(\alpha)|_v = \prod_{i < j} |\alpha_i - \alpha_j|_v^2$, where α_i, α_j are the roots of f_α . Consequently, for any point in the support of $\mathbf{1}_\delta$, we have

$$\prod_{i=1}^n \prod_{j \neq i} |\alpha_i - \alpha_j|_v \leq \delta.$$

The pigeonhole principle then implies that there must be some i for which

$$\prod_{j \neq i} |\alpha_i - \alpha_j|_v \leq \delta^{1/n}.$$

If we let $\alpha'_j = \alpha_i - \alpha_j$ for each $j \neq i$, it thus follows that $\prod_{j \neq i} |\alpha'_j|_v \leq \delta^{1/n}$. Since \mathcal{O}_{K_v} is compact, the measure of points satisfying this condition is $O_n(\delta^{1/n})$, independent of α_i . We therefore conclude in this case that

$$\int_{\mathcal{O}_{K_v}} \mathbf{1}_\delta(f_\alpha) d\mu(\alpha) \ll_n \delta^{\frac{1}{n}}.$$

We now turn to the case that K_v is not the totally split algebra. The idea in this case is substantially the same, but requires the notation from Section 3. The discriminant $\text{Disc}(\alpha)$ satisfies

$$|\text{Disc}(\alpha)|_v = \prod_{i=1}^n \prod_{j \neq i} |\lambda_i - \lambda_j|_v.$$

Proceeding as in the totally split case, we conclude that for α in the support of $\mathbf{1}_\delta$, there must be some i for which

$$\prod_{j \neq i} |\lambda_i - \lambda_j|_v \leq \delta^{\frac{1}{n}}. \quad (4.2)$$

Accounting for the determinant of the matrix M_{K_v} , sending the coordinates \mathbf{a} to the roots $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n)$, we conclude that

$$\int_{\mathcal{O}_{K_v}} \mathbf{1}_\delta(f_\alpha) d\mu(\alpha) \ll_n |\text{Disc}(K_v)|_v^{-1/2} \delta^{\frac{1}{n}}.$$

We therefore find that

$$\delta^{1/2} \sum_{[K_v:\mathbb{Q}_v]=n} \frac{|\text{Disc}(K_v)|_v^{1/2}}{|\text{Aut}(K_v)|} \int_{\mathcal{O}_{K_v}} \mathbf{1}_\delta(f_\alpha) d\mu(\alpha) \ll_n \delta^{\frac{1}{2}+\frac{1}{n}} \sum_{[K_v:\mathbb{Q}_v]=n} \frac{1}{|\text{Aut}(K_v)|} \ll_n \delta^{\frac{1}{2}+\frac{1}{n}},$$

since the summation may be bounded in terms of the number of étale algebras of degree n , which in turn may be bounded solely in terms of n . \square

Remark 4.2. The bound (4.1) is sharp for an infinite sequence of δ tending to 0. In particular, one can realize a lower bound of magnitude $\gg \delta^{\frac{1}{2}+\frac{1}{n}}$ for all δ when v is an archimedean valuation. One can also realize this lower bound for the nonarchimedean totally split étale algebras $K_v \simeq \mathbb{Q}_v^n$ and $\delta = p^{-k}$ when $k = mn(n-1)$ for some k and $m \geq 1$ since the product in (4.2) has $n(n-1)$ factors in these cases, and therefore can be realized in these cases. For many other δ , the bound can be improved using the discrete nature of the valuation group. In our optimization process below, we have no freedom in choosing δ and cannot exploit such improvements. Therefore, we rely on the upper bound (4.1).

We now use Proposition 4.1 to bound the number of integral polynomials with small archimedean discriminant.

Lemma 4.3. *Let $n \geq 3$, $Y \geq 1$, and H sufficiently large in terms of n . Then the number of polynomials $f(x) \in \mathbb{Z}[x]$ of the form $f(x) = x^n + c_1x^{n-1} + \dots + c_n$ with $|c_i| \leq H^i$ and $\text{disc}(f) \leq H^{n^2-n}/Y$ is $O_n(H^{\frac{n^2+n}{2}}/Y^{\frac{1}{2}+\frac{1}{n}} + H^{\frac{n^2+n}{2}-1})$.*

The key idea behind Lemma 4.3 is that the coefficients of the polynomials $f(x)$ with small discriminant must lie in a compact region in \mathbb{R}^n with small volume, and thus we should expect there to be relatively few polynomials whose coefficients lie in this region. This is made rigorous by means of the Lipschitz principle from [Dav51]:

Lemma 4.4 (Davenport). *If $\Omega \subseteq \mathbb{R}^n$ is a compact semialgebraic region (i.e., cut out by algebraic inequalities), then the number of lattice points in the intersection satisfies the bound*

$$\mathbb{Z}^n \cap \Omega = \text{vol}(\Omega) + O(\max_{\pi} \text{vol}(\pi(\Omega))), \quad (4.3)$$

where the maximum runs over the projections π of \mathbb{R}^n onto its various coordinate hyperplanes (i.e., the regions in \mathbb{R}^d for $d < n$ obtained by “forgetting” $n-d$ of the coordinates). The implicit constants depend only on the dimension n and the degrees of the equations defining Ω .

Proof of Lemma 4.3. Define $\Omega_{H,Y} \subseteq \mathbb{R}^n$ to be

$$\Omega_{H,Y} := \{(c_1, \dots, c_n) \in \mathbb{R}^n : |c_i| \leq H^i, \text{disc}(x^n + c_1x^{n-1} + \dots + c_n) \leq H^{n^2-n}/Y\}.$$

The maximum volume of the coordinate projections of $\Omega_{H,Y}$ is trivially $O_n(H^{\frac{n^2+n}{2}-1})$, coming from forgetting the discriminant condition and using the coordinate projection $(c_1, \dots, c_n) \mapsto (c_2, \dots, c_n)$. By Lemma 4.4, the statement thus reduces to showing that $\text{vol}(\Omega_{H,Y}) \ll_n H^{\frac{n^2+n}{2}} / Y^{\frac{1}{2} + \frac{1}{n}}$. Since the region $\Omega_{H,Y}$ is obtained by an anisotropic dilation we find $\text{vol}(\Omega_{H,Y}) = H^{\frac{n^2+n}{2}} \text{vol}(\Omega_{1,Y})$. Thus it suffices to bound $\text{vol}(\Omega_{1,Y})$. Taking $\delta = Y^{-1}$ and $\mathbb{Q}_v = \mathbb{R}$ in Proposition 4.1, we immediately deduce that

$$\text{vol}(\Omega_{1,Y}) \ll_n Y^{-\frac{1}{2} - \frac{1}{n}}. \quad (4.4)$$

This completes the proof. \square

This lemma implies Corollary 2.5, used in §2.1 to prove our main theorem.

Proof of Corollary 2.5. Take $Y = H^2$ in Lemma 4.3. \square

5 Nonarchimedean Fourier transforms

Recall that if $f(x)$ is irreducible and cuts out the field K , then its discriminant satisfies $\text{disc}(f) = \text{Disc}(K)[\mathcal{O}_K : \mathbb{Z}[\alpha]]^2$, where α is a root of f over K , and where $[\mathcal{O}_K : \mathbb{Z}[\alpha]] =: \text{index}(f)$. In this section, we assemble the main technical ingredients that will be used in the proof of Theorem 2.6. This theorem bounds the number of integral polynomials f whose index is large, but for which $\text{rad}(\text{index}(f))$ is small. This condition implies that such polynomials will have discriminants divisible by large powers of primes. We therefore begin with the following simple lemma.

Lemma 5.1. *Let $n \geq 2$ and $k \geq 1$, and let p be prime. The set of monic polynomials $f \in \mathbb{Z}[x]$ for which p^{2k} divides the discriminant of f is determined by congruence conditions $(\text{mod } p^{2k})$ with relative density $\ll p^{-k - \frac{2k}{n}}$.*

Proof. The fact that this condition is determined from congruence conditions follows from the fact that the discriminant of a polynomial is a polynomial in its coefficients. The density of these congruence conditions may be determined by computing the Haar measure of the set of monic polynomials in $\mathbb{Z}_p[x]$ satisfying this condition. Therefore, the claim about the density follows from Proposition 4.1 by taking $\mathbb{Q}_v = \mathbb{Q}_p$ and $\delta = p^{-2k}$. \square

We now introduce the main object of study in this section. Let $\mathcal{P}_n(\mathbb{Z}/p^{2k}\mathbb{Z})$ be the set of monic degree n polynomials over $\mathbb{Z}/p^{2k}\mathbb{Z}$, and let $\psi_{p^{2k}}$ be the characteristic function of the subset of polynomials whose discriminant is congruent to 0 $(\text{mod } p^{2k})$. For any $\mathbf{u} \in \mathbb{Z}^n$, we define the Fourier transform $\widehat{\psi_{p^{2k}}}(\mathbf{u})$ of $\psi_{p^{2k}}$ by

$$\widehat{\psi_{p^{2k}}}(\mathbf{u}) := \frac{1}{p^{2kn}} \sum_{\mathbf{c} \in (\mathbb{Z}/p^{2k}\mathbb{Z})^n} \psi_{p^{2k}}(f_{\mathbf{c}}) \exp\left(\frac{2\pi i \langle \mathbf{c}, \mathbf{u} \rangle}{p^{2k}}\right), \quad (5.1)$$

where $f_{\mathbf{c}}(x) = x^n + c_1 x^{n-1} + \dots + c_n$ and $\langle \cdot, \cdot \rangle$ denotes the standard inner product.

We begin with a series of lemmas that will be used to determine the support of $\widehat{\psi_{p^{2k}}}$.

Lemma 5.2. Write the monic polynomial f as $f(x) = x^n - \sigma_1 x^{n-1} + \cdots + (-1)^k \sigma_k x^{n-k} + \cdots + (-1)^n \sigma_n$. For each $1 \leq i \leq n$, define $D_i := \frac{\partial \text{disc} f}{\partial \sigma_i} \in \mathbb{Z}[\sigma_1, \dots, \sigma_n]$. Then the discriminant $\text{disc}(f)$ and partial derivatives D_i satisfy

1. For any $1 \leq r \leq r+k \leq n$ and $1 \leq s-k \leq s \leq n$, we have $\text{disc}(f) \mid (D_r D_s - D_{r+k} D_{s-k})$.
2. $\sum_{1 \leq i \leq n} D_i \cdot (n+1-i) \cdot \sigma_{i-1} = 0$ where we define $\sigma_0 = 1$.

The second part of the above lemma can be derived from equation (1.27) in Chapter 12 of [GKZ08], but we were unaware of this reference until recently, so we provide a proof.

Proof. Our first goal is to give an expression for $D_i(\boldsymbol{\sigma}(\boldsymbol{\lambda})) \in \mathbb{C}[\boldsymbol{\lambda}]$. Recall the change of variables to go from roots $\boldsymbol{\lambda}$ to coefficients $\boldsymbol{\sigma}(\boldsymbol{\lambda})$ of a polynomial described in Section 3. The Jacobian matrix for this change of variables is $B(\boldsymbol{\lambda})$ where $B_{ij} = \frac{\partial c_i}{\partial \lambda_j}$. A quick computation verifies that

$$\frac{\partial c_i}{\partial \lambda_j} = c_{i-1}(\lambda_1, \dots, \hat{\lambda}_j, \dots, \lambda_n),$$

where we write $\hat{\lambda}_j$ to indicate that the j -th coordinate should be omitted. Below, we use the shorthand $\hat{\boldsymbol{\lambda}}_j = (\lambda_1, \dots, \hat{\lambda}_j, \dots, \lambda_n)$. When $\lambda_i \neq \lambda_j$ for all $i \neq j$, we define the matrix $A(\boldsymbol{\lambda}) = (A_{ij})$ by

$$A_{ij} = \lambda_i^{n-j} (-1)^{j+1} / \prod_{k \neq i} (\lambda_i - \lambda_k).$$

Then one can verify that

$$(AB)_{ij} \prod_{k \neq i} (\lambda_i - \lambda_k) = (\lambda_i^{n-1} \sigma_0(\hat{\boldsymbol{\lambda}}_j) - \lambda_i^{n-2} \sigma_1(\hat{\boldsymbol{\lambda}}_j) + \cdots + (-1)^{n-1} \sigma_{n-1}(\hat{\boldsymbol{\lambda}}_j)) = \prod_{k \neq j} (\lambda_i - \lambda_k).$$

To see the last equality, we used the identity

$$\prod_{i=1}^{n-1} (x - a_i) = \sum_{k=0}^{n-1} (-1)^k \sigma_k(\mathbf{a}) x^{n-1-k} \quad (5.2)$$

for expressing a polynomial with roots $\mathbf{a} = (a_1, \dots, a_{n-1})$ in terms of the elementary symmetric polynomials in these roots. This shows that $(AB)_{ij}$ is 0 if $i \neq j$ and is equal to 1 if $i = j$, hence $AB = I$.

By the inverse function theorem and chain rule, we obtain the inverse function $\lambda_i = \lambda_i(\boldsymbol{\sigma})$ with $\frac{\partial \lambda_i}{\partial \sigma_j} = A_{ij}$, and $\frac{\partial D}{\partial \sigma_i} = \sum_j \frac{\partial D}{\partial \lambda_j} \frac{\partial \lambda_j}{\partial \sigma_i} = \sum_j \frac{\partial D}{\partial \lambda_j} A_{ji}$.

By an abuse of notation, we write $\text{disc}(\mathbf{c}) := D(\mathbf{c}) := \text{disc}(f_{\mathbf{c}})$, and with $f = x^n - \sigma_1 x^{n-1} + \cdots + (-1)^n \sigma_n$ as in the lemma statement, we write $D = \text{disc}(f)$. For $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_n) = \boldsymbol{\sigma}(\boldsymbol{\lambda})$, we note that (5.2) also shows that the roots of f are precisely the λ_i . Thus $D(\boldsymbol{\sigma}(\boldsymbol{\lambda})) = \prod_{i < j} (\lambda_i - \lambda_j)^2$, and we have $\frac{\partial D}{\partial \lambda_j} = 2D \sum_{k \neq j} (\lambda_j - \lambda_k)^{-1}$. This implies that

$$D_i := \frac{\partial D}{\partial \sigma_i} = 2D \sum_j \lambda_j^{n-i} (-1)^{i+1} \cdot \prod_{k \neq j} (\lambda_j - \lambda_k)^{-1} \cdot \left(\sum_{k \neq j} (\lambda_j - \lambda_k)^{-1} \right) = 2D \sum_j \lambda_j^{n-i} (-1)^{i+1} \cdot W_j,$$

where we denote

$$W_j = \prod_{k \neq j} (\lambda_j - \lambda_k)^{-1} \cdot \sum_{k \neq j} (\lambda_j - \lambda_k)^{-1}.$$

Observe that

$$DW_j = \prod_{i < k, i \neq j, k \neq j} (\lambda_i - \lambda_k)^2 \cdot \prod_{k \neq j} (\lambda_j - \lambda_k) \cdot \sum_{k \neq j} (\lambda_j - \lambda_k)^{-1}$$

is a polynomial in $\mathbb{C}[\lambda]$. We thus obtain $D_i(\lambda)$ as a polynomial in the Zariski open set $D(\lambda) \neq 0$ in the affine space of λ . As $D_i(\sigma(\lambda))$ agrees with $D_i(\lambda)$ in that open set, they agree everywhere in λ affine space (and not merely when $\lambda_i \neq \lambda_j$ for all $i \neq j$).

We now evaluate D_i . Notice that DW_j is zero if λ_j is a triple root of f , or if there exists a double root $\lambda_i = \lambda_k$ with $i, k \neq j$. If there is a unique double root $\lambda_j = \lambda_{k_0}$, then $DW_j = \prod_{i < k, i \neq j, k \neq j} (\lambda_i - \lambda_k)^2 \prod_{k \neq j, k_0} (\lambda_j - \lambda_k) = DW_{k_0}$. Stated differently — if f has two distinct pairs of double roots or a triple root, then $D_i = 0$ for all i . If there is a unique double root λ_j , then for any i we compute

$$D_i = (-1)^{i+1} \lambda_j^{n-i} \cdot 4DW_j.$$

Therefore as a polynomial in $\mathbb{C}[\sigma]$, it follows that

$$D_r D_s - D_{r+k} D_{s-k} \equiv (4DW_j)^2 (-1)^{r+s} \cdot (\lambda_j^{2n-r-s} - \lambda_j^{2n-r-s}) = 0$$

on the variety cut out by $D(\sigma) = 0$. Since $D(\sigma)$ is irreducible, it follows from Hilbert's Nullstellensatz that $D \mid D_r D_s - D_{r+k} D_{s-k}$.

In order to prove the second statement, notice that discriminant is translation invariant. That is, $\text{disc}(\lambda_1 + t, \lambda_2 + t, \dots, \lambda_n + t) = \text{disc}(\lambda_1, \lambda_2, \dots, \lambda_n)$. Now we can consider the function $\sigma_i(t) := \sigma_i(\lambda_1 + t, \lambda_2 + t, \dots, \lambda_n + t)$. It is clear that $\frac{d\sigma_i(t)}{dt} = \sigma_{i-1}(\lambda_1 + t, \dots, \lambda_n + t)(n - i + 1)$. Therefore

$$\left. \frac{dD(\sigma(t))}{dt} \right|_{t=0} = \sum_i \frac{\partial D}{\partial \sigma_i} \frac{d\sigma_i}{dt} \Big|_{t=0} = \sum_i D_i \cdot \sigma_{i-1} \cdot (n - i + 1) = 0,$$

completing the proof. \square

This lemma is general. We exploit the relationship $D \mid D_r D_s - D_{r+k} D_{s-k}$ below, but we note that there are many other algebraic relationships on the discriminant and its derivatives that might yield further refinements on the structure of the support of $\psi_{p^{2k}}$. In practice, we use the following more specific lemma.

Lemma 5.3. *Given $\mathbf{c} \in \mathbb{Z}^n$ corresponding to a polynomial $f_{\mathbf{c}}(x)$ with $\text{disc}(f_{\mathbf{c}}) \equiv 0 \pmod{p^{2k}}$, let $\mathbf{D}_{\mathbf{c}} := (\frac{\partial \text{disc}(f)}{\partial c_1}, \dots, \frac{\partial \text{disc}(f)}{\partial c_n})$ denote the gradient vector of the discriminant function, and let $v_i(\mathbf{c})$ denote the valuation at p of $\frac{\partial \text{disc}(f)}{\partial c_i}(\mathbf{c}) = D_i(\mathbf{c})$.*

Then there either exists $a \in \mathbb{Z}_{\geq 0}$ such that

$$\min\{v_i(\mathbf{c}), k\} = \min\{v_n(\mathbf{c}) + (n - i)a, k\} \tag{5.3}$$

or $b \in \mathbb{Z}$ with $0 \leq b \leq \min(\text{val}_p(n), k)$ such that

$$\min\{v_i(\mathbf{c}), k\} = \min\{v_1(\mathbf{c}) + (i - 1)b, k\}. \tag{5.4}$$

Intuitively, $\min\{v_i(\mathbf{c}), k\}$ almost forms an arithmetic progression, except that when a term is greater than k we change it to k . Thus this lemma shows that the support of $\psi_{p^{2k}}$ is constrained to these “near arithmetic progressions”.

Proof. By Lemma 5.2 (1), we see that when $p^{2k} \mid \text{disc}(f_{\mathbf{c}})$, $p^{2k} \mid D_r(\mathbf{c})D_s(\mathbf{c}) - D_{r+\ell}(\mathbf{c})D_{s-\ell}(\mathbf{c})$ for all relevant r, s, ℓ . This implies the valuation relations $\min\{v_r(\mathbf{c}) + v_s(\mathbf{c}), 2k\} = \min\{v_{r+\ell}(\mathbf{c}) + v_{s-\ell}(\mathbf{c}), 2k\}$. Specializing to $r = s = i$ and $\ell = 1$ gives that $\min\{2v_i(\mathbf{c}), 2k\} = \min\{v_{i-1}(\mathbf{c}) + v_{i+1}(\mathbf{c}), 2k\}$. Thus $(v_{i-1}(\mathbf{c}), v_i(\mathbf{c}), v_{i+1}(\mathbf{c}))$ is an arithmetic progression when $v_i(\mathbf{c}) < k$ and “nearly” an arithmetic progression otherwise.

It is not possible for the sequence $\min\{v_i(\mathbf{c}), k\}$ to initially decrease and later decrease. For otherwise, there would exist r and s , with $r < s - 1$, such that $\min\{v_r(\mathbf{c}), k\}$ and $\min\{v_s(\mathbf{c}), k\}$ are smaller than $\min\{v_i(\mathbf{c}), k\}$ for each $r < i < s$. But this would contradict the valuation relation above with $\ell = 1$.

Hence the sequence $\min\{v_i(\mathbf{c}), k\}$ is either non-increasing (giving (5.3)) or non-decreasing (giving (5.4)). Note that Lemma 5.2(2) implies that $nD_1(\mathbf{c})$ is in the ideal generated by $(D_2(\mathbf{c}), \dots, D_n(\mathbf{c}))$, which restricts $b \leq \min\{\text{val}_p(n), k\} \leq n$ in (5.4). \square

Lemma 5.3 implies that the support of $\widehat{\psi_{p^{2k}}}$ is also on “near arithmetic progressions.”

Lemma 5.4 (Support on near arithmetic progressions). *For $\mathbf{u} = (u_1, \dots, u_n) \in (\mathbb{Z}/p^{2k}\mathbb{Z})^n$, we have that*

$$\widehat{\psi_{p^{2k}}}(\mathbf{u}) = 0$$

unless \mathbf{u} satisfies one of the two “near arithmetic progression” properties that

$$\min\{v_p(u_i), k\} = \min\{v_p(u_n) + (n-i)a, k\} \quad (5.5)$$

for some $a \in \mathbb{Z}_{\geq 0}$, or

$$\min\{v_p(u_i), k\} = \min\{v_p(u_1) + (i-1)b, k\} \quad (5.6)$$

for some $b \in \mathbb{Z}$ with $0 \leq b \leq \min\{v_p(n), k\}$.

Proof. For each $\mathbf{c} \in (\mathbb{Z}/p^{2k}\mathbb{Z})^n$ such that $p^{2k} \mid \text{disc}(f_{\mathbf{c}})$, we associate the locus

$$P_{\mathbf{c}} := \{\mathbf{v} \in (\mathbb{Z}/p^{2k}\mathbb{Z})^n : p^{2k} \mid \text{disc}(f_{\mathbf{v}}), \mathbf{v} - \mathbf{c} \in p^k \cdot (\mathbb{Z}/p^{2k}\mathbb{Z})^n\}.$$

Two such $P_{\mathbf{c}}$ and $P_{\mathbf{c}'}$ are equal if and only if $\mathbf{c} - \mathbf{c}' \equiv 0 \pmod{p^k}$. Thus we can decompose the set of congruence classes of polynomials whose discriminant is divisible by p^{2k} into a disjoint union of $P_{\mathbf{c}}$ over a set C of representatives \mathbf{c} from $p^k(\mathbb{Z}/p^{2k}\mathbb{Z})^n$ satisfying $p^{2k} \mid \text{disc}(f_{\mathbf{c}})$, giving

$$\widehat{\psi_{p^{2k}}} = \sum_{\mathbf{c} \in C} \widehat{\mathbf{1}_{P_{\mathbf{c}}}}.$$

This leads us to study $\widehat{\mathbf{1}_{P_{\mathbf{c}}}}$. Below, we use the notation $\mathbf{D}_{\mathbf{c}}$ for the gradient vector of the discriminant function and the notation v_i for the p -valuation for the i -th coordinate of $\mathbf{D}_{\mathbf{c}}$ as in Lemma 5.3. From the Taylor expansion

$$\text{disc}(f_{\mathbf{v}}) \equiv \text{disc}(f_{\mathbf{c}}) + \mathbf{D}_{\mathbf{c}} \cdot (\mathbf{v} - \mathbf{c}) \pmod{p^{2k}}$$

and the fact that $p^{2k} \mid \text{disc}(f_{\mathbf{c}})$, we see that $P_{\mathbf{c}}$ can be written as

$$P_{\mathbf{c}} = \left\{ \mathbf{v} \in (\mathbb{Z}/p^{2k}\mathbb{Z})^n : \mathbf{D}_{\mathbf{c}} \cdot \frac{\mathbf{v} - \mathbf{c}}{p^k} \equiv 0 \pmod{p^k} \right\}. \quad (5.7)$$

Direct computation on the definition shows

$$\widehat{\mathbf{1}_{P_{\mathbf{c}}}}(\boldsymbol{\xi}) = \frac{1}{p^{2kn}} \exp\left(2\pi i \frac{\langle \boldsymbol{\xi}, \mathbf{c} \rangle}{p^{2k}}\right) \sum_{\mathbf{v} \in P_{\mathbf{c}}} \exp\left(2\pi i \frac{\langle \boldsymbol{\xi}, \mathbf{v} - \mathbf{c} \rangle}{p^{2k}}\right). \quad (5.8)$$

Let $w := \min(v_i(\mathbf{c}), k)$ denote the minimum valuation among the coordinates of $\mathbf{D}_{\mathbf{c}}$. We will now show that

$$\widehat{\mathbf{1}_{P_{\mathbf{c}}}}(\boldsymbol{\xi}) = \begin{cases} \exp(2\pi i \frac{\langle \boldsymbol{\xi}, \mathbf{c} \rangle}{p^{2k}}) \cdot \frac{p^{w-k}}{p^{kn}}, & \boldsymbol{\xi} \equiv \alpha \mathbf{D}_{\mathbf{c}} \pmod{p^k} \text{ for some } \alpha \in \mathbb{Z}/p^k\mathbb{Z} \\ 0 & \text{otherwise.} \end{cases} \quad (5.9)$$

To see this, first note that by (5.7), the set $P_{\mathbf{c}} - \mathbf{c}$ forms an additive group. This implies that

$$\sum_{\mathbf{v} \in P_{\mathbf{c}}} \exp\left(2\pi i \frac{\langle \boldsymbol{\xi}, \mathbf{v} - \mathbf{c} \rangle}{p^{2k}}\right) = \exp\left(2\pi i \frac{\langle \boldsymbol{\xi}, \mathbf{u} - \mathbf{c} \rangle}{p^{2k}}\right) \sum_{\mathbf{v} \in P_{\mathbf{c}}} \exp\left(2\pi i \frac{\langle \boldsymbol{\xi}, \mathbf{v} - \mathbf{c} \rangle}{p^{2k}}\right) \quad (5.10)$$

for any $\mathbf{u} \in P_{\mathbf{c}}$.

When $\boldsymbol{\xi} \equiv \alpha \mathbf{D}_{\mathbf{c}} \pmod{p^k}$ for some $\alpha \in \mathbb{Z}/p^k\mathbb{Z}$, all summands in (5.8) are equal to 1 by the definition of $P_{\mathbf{c}}$ in (5.7). The first case of (5.9) follows from the fact that $|P_{\mathbf{c}}| = p^{k(n-1)+w}$. If $\boldsymbol{\xi}$ cannot be written as $\alpha \mathbf{D}_{\mathbf{c}}$, then there must be some $\mathbf{u} \in P_{\mathbf{c}}$ such that $\langle \boldsymbol{\xi}, \mathbf{u} - \mathbf{c} \rangle \not\equiv 0 \pmod{p^{2k}}$. Inserting this choice of \mathbf{u} into (5.10) shows that $\widehat{\mathbf{1}_{P_{\mathbf{c}}}}(\boldsymbol{\xi}) = 0$. Thus (5.9) is shown.

It follows that for all \mathbf{u} in the support of $\widehat{\psi}_{p^{2k}}$, $\mathbf{u} \equiv \alpha \mathbf{D}_{\mathbf{c}}$ for some \mathbf{c} . The near arithmetic progression conditions (5.5) and (5.6) are then implied by Lemma 5.3. \square

For the phases $\mathbf{u} = (u_1, \dots, u_n)$ for which the Fourier transform $\widehat{\psi}_{p^{2k}}(\mathbf{u})$ does not vanish, we will need an improvement over the trivial bound $|\widehat{\psi}_{p^{2k}}(\mathbf{u})| \leq |\widehat{\psi}_{p^{2k}}(\mathbf{0})| \ll_n p^{-k-\frac{2k}{n}}$ given by Lemma 5.1. In fact, for our purposes, it will suffice to restrict our attention to those phases for which only u_1 and u_2 are possibly non-zero. For the phases $\mathbf{u} = (u_1, \dots, u_n)$ in which only u_1 is non-zero, we observe that the Fourier transform is typically 0.

Lemma 5.5. *Let p be prime, $n \geq 6$, and $k \geq 3$. Suppose that $u_1 \in \mathbb{Z}$, and define $\mathbf{u} \in \mathbb{Z}^n$ by $\mathbf{u} = (u_1, 0, 0, \dots, 0)$. Then $\widehat{\psi}_{p^{2k}}(\mathbf{u}) = 0$ if $p^{2k} / \gcd(p^{2k}, n)$ does not divide u_1 .*

Proof. Since only $u_1 \neq 0$, by definition, we have

$$\widehat{\psi}_{p^{2k}}(\mathbf{u}) = \frac{1}{p^{2nk}} \sum_{\mathbf{c} \in (\mathbb{Z}/p^{2k}\mathbb{Z})^n} \psi_{p^{2k}}(f_{\mathbf{c}}) \exp\left(\frac{2\pi i c_1 u_1}{p^{2k}}\right).$$

As the map $x \mapsto x+1$ induces a bijection from the support of $\psi_{p^{2k}}$ to itself (that is, on the set of polynomials whose discriminant is divisible by p^{2k}), sending $c_1 = c_1(f)$ to $c_1 + n$ for each f , we also have

$$\widehat{\psi}_{p^{2k}}(\mathbf{u}) = \frac{1}{p^{2nk}} \sum_{\mathbf{c} \in (\mathbb{Z}/p^{2k}\mathbb{Z})^n} \psi_{p^{2k}}(f_{\mathbf{c}}) \exp\left(\frac{2\pi i (c_1 + n) u_1}{p^{2k}}\right).$$

This implies that either $\widehat{\psi}_{p^{2k}}(\mathbf{u}) = 0$ or $p^{2k} \mid nu_1$. \square

To study phases $\mathbf{u} = (u_1, \dots, u_n)$ with both u_1 and u_2 non-zero, we'll use the following additional result.

Lemma 5.6. *Let p be prime, $n \geq 6$, and $k \geq 3$. Suppose that $u_1, u_2 \in \mathbb{Z}$, and define $\mathbf{u} \in \mathbb{Z}^n$ by $\mathbf{u} = (u_1, u_2, 0, \dots, 0)$. Then*

$$\widehat{\psi_{p^{2k}}(\mathbf{u})} \ll_n p^{-\frac{2nk}{3}} \gcd(u_2, p^{2k})^{\frac{n}{3}}.$$

Proof. Let $e_{p^{2k}}: \mathbb{Z}_p \rightarrow \mathbb{C}$ be the continuous extension of the map defined on integers x by $e_{p^{2k}}(x) := \exp(\frac{2\pi i x}{p^{2k}})$. By definition, it follows that we may rewrite the Fourier transform as

$$\widehat{\psi_{p^{2k}}(\mathbf{u})} = \int_{\mathcal{P}_n(\mathbb{Z}_p)} \psi_{p^{2k}}(f) e_{p^{2k}}(u_1 c_1(f) + u_2 c_2(f)) d\nu(f),$$

where $c_1(f)$ and $c_2(f)$ respectively denote the coefficients of x^{n-1} and x^{n-2} in f . Applying the change of variables from Section 3, we find this integral is equal to

$$\sum_{[K_p: \mathbb{Q}_p]=n} \frac{|\text{Disc}(K_p)|_p^{1/2}}{|\text{Aut}(K_p)|} \int_{\mathcal{O}_{K_p}} |\text{Disc}(f_\alpha)|_p^{1/2} \psi_{p^{2k}}(f_\alpha) e_{p^{2k}}(-u_1 \sigma_1(\boldsymbol{\lambda}) + u_2 \sigma_2(\boldsymbol{\lambda})) d\mu(\alpha), \quad (5.11)$$

where f_α is the characteristic polynomial of α and $\sigma_i(\boldsymbol{\lambda})$ denotes the i -th elementary symmetric function in the roots $\boldsymbol{\lambda}$. Fixing the étale algebra $K_p = F_1 \times \dots \times F_r$ (with $n_i := \deg(F_i)$), we note that σ_1 is linear as a polynomial in $\boldsymbol{\lambda}$, and is therefore also linear in the choice of coordinates $\mathbf{a} \in \mathbb{Z}_p^n$ from Section 3 and the proof of Proposition 4.1 (which we use in this proof as well). Additionally, σ_2 is a quadratic form in $\boldsymbol{\lambda}$ with Gram matrix

$$Q = \begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ \vdots & & \ddots & & \vdots \\ 1 & 1 & 1 & \dots & 0 \end{pmatrix},$$

that is, $\sigma_2(\boldsymbol{\lambda}) = \frac{1}{2} \boldsymbol{\lambda}^T Q \boldsymbol{\lambda}$, where we regard $\boldsymbol{\lambda}$ as a column vector. As $\boldsymbol{\lambda} = M_{K_p} \mathbf{a}$, we therefore find that σ_2 , as a polynomial in \mathbf{a} , has Gram matrix $M_{K_p}^T Q M_{K_p}$. For convenience in what is to come, we note now that $\det Q = (-1)^{n-1} (n-1)$. Thus $|\det M_{K_p}^T Q M_{K_p}|_p \ll_n |\text{Disc}(K_p)|_p$, where M_{K_p} is as in Section 3.

Similarly, using the formula in Section 3, if one considers σ_2 (as a polynomial in \mathbf{a}) as a quadratic form modulo p , this computation shows that we have $|\text{Disc}(K_p)|_p \ll p^{-\text{corank}_{\mathbb{F}_p}(\sigma_2)}$, as the corank represents the number of column vectors that are linearly dependent p -adically.

For each $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{Z}_p^n$ associated to a point $\alpha \in \mathcal{O}_{K_p}$ in the support of $\psi_{p^{2k}}$, we introduce the following quantities:

- We set $\ell = k - \frac{v_p(u_2)}{2}$, and in the following we will write $a \equiv b \pmod{p^\ell}$ to mean that $v_p(a-b) \geq \ell$, including the case when ℓ is a half integer.
- We write S for the set of i (or, by abuse of notation, the set of a_i) with the following property: one of the roots for which a_i is a coefficient is subject to a congruence $\lambda \equiv \lambda' \pmod{p^\ell}$, where λ' is some other root, and where the congruence is as polynomials in the appropriate integral bases.

- We write $m := |S|$.

If $a_i \in S$, then using the Galois action we see that *every* root for which a_i is a coefficient will be subject to such a congruence, and thus we may regard this congruence as being of the form $\alpha \equiv \gamma(\alpha') \pmod{p^\ell}$ for some nontrivial $\gamma \in \text{Gal}(\overline{\mathbb{Q}_p}/\mathbb{Q}_p)$.

We distinguish two cases:

1. If $\alpha_j \equiv \gamma(\alpha_j) \pmod{p^\ell}$ for some nontrivial $\gamma \in \text{Gal}(\widetilde{F_j}/\mathbb{Q}_p)$, let $|\gamma|$ denote the order of γ . Taking traces shows that $|\gamma|\alpha_j$ must be congruent $\pmod{p^\ell}$ to an element of the subfield F_j^γ of F_j fixed by γ . Since γ acts nontrivially on the conjugates of F_j , this subfield has degree at most $n_j/2$, from which it follows that at least $n_j/2$ values r_j are greater than ℓ . This will be satisfied for a proportion of $O(p^{-\ell n_j/2})$ of the possible α_j .
2. We divide the set of remaining α into equivalence classes where $\alpha_{j_1} \equiv \gamma_2(\alpha_{j_2}) \equiv \dots \equiv \gamma_t(\alpha_{j_t}) \pmod{p^\ell}$ for some $t \geq 2$. Once the γ are fixed, any one of the α_j determines all the rest $\pmod{p^\ell}$. Assuming without loss of generality that α_{j_1} is of minimal degree, these congruences will be satisfied for a proportion of $O(\min(p^{-\ell(n_{j_2}+\dots+n_{j_t})}))$ of the possible tuples $(\alpha_{j_1}, \dots, \alpha_{j_t})$.

There are $O_n(1)$ ways of partitioning the set of α into equivalence classes as above and choosing the γ . Altogether we obtain a density $\ll_n p^{-\ell N/2}$, where N is the total of the degrees n_j over those α_j corresponding to at least one coordinate in S . As $N \geq m = |S|$, this density is $\ll_n p^{-\ell m/2}$. Moreover, the above procedure identifies $\geq m/2$ pairs of roots, distinct but not necessarily disjoint, which are congruent $\pmod{p^\ell}$, so we have

$$|\text{Disc}(f_\alpha)|_p^{1/2} \leq p^{-\frac{m\ell}{2}} |\text{Disc}(K_p)|_p^{1/2} \quad (5.12)$$

for any such α . So far we obtain a bound

$$\ll p^{-m\ell} |\text{Disc}(K_p)|_p^{1/2} \quad (5.13)$$

for the integral in (5.11).

For the $n - m$ remaining coordinates $a_i \notin S$, we may change a_i by an arbitrary multiple of $p^{\lceil \ell \rceil}$ while preserving the discriminant. Motivated by the method of (non)stationary phase, by Taylor expanding, we then observe that

$$u_2 \sigma_2(\mathbf{a} + b_i p^{\lceil \ell \rceil}) - u_1 \sigma_1(\mathbf{a} + b_i p^{\lceil \ell \rceil}) \equiv u_2 \sigma_2(\mathbf{a}) - u_1 \sigma_1(\mathbf{a}) + b_i p^{\lceil \ell \rceil} \cdot \left(u_2 \frac{\partial \sigma_2}{\partial a_i}(\mathbf{a}) - u_1 \frac{\partial \sigma_1}{\partial a_i}(\mathbf{a}) \right) \pmod{p^{2k}},$$

where we use b_i to indicate a vector whose only nonzero component is b_i in the i -th coordinate. We have used that $\sigma_1(\mathbf{a})$ is linear in a_i and the relationship between ℓ and u_2 in omitting higher derivatives. This implies that the integral

$$\int_{b_i \in \mathbb{Z}_p} |\text{Disc}(f_\alpha)|_p^{1/2} \psi_{p^{2k}}(f_\alpha) e_{p^{2k}}(-u_1 \sigma_1(\mathbf{a} + b_i p^{\lceil \ell \rceil}) + u_2 \sigma_2(\mathbf{a} + b_i p^{\lceil \ell \rceil})) db_i$$

will vanish unless the congruence $u_2 \frac{\partial \sigma_2}{\partial a_i}(\mathbf{a}) - u_1 \frac{\partial \sigma_1}{\partial a_i}(\mathbf{a}) \equiv 0 \pmod{p^{2k - \lceil \ell \rceil}}$ is satisfied. Consequently, we may restrict the integral

$$\int_{\mathcal{O}_{K_p}} |\text{Disc}(f_\alpha)|_p^{1/2} \psi_{p^{2k}}(f_\alpha) e_{p^{2k}}(-u_1 \sigma_1(\boldsymbol{\lambda}) + u_2 \sigma_2(\boldsymbol{\lambda})) d\mu(\alpha) \quad (5.14)$$

to those points satisfying the congruence $u_2 \frac{\partial \sigma_2}{\partial a_i}(\mathbf{a}) - u_1 \frac{\partial \sigma_1}{\partial a_i}(\mathbf{a}) \equiv 0 \pmod{p^{2k-\lceil \ell \rceil}}$ for each $a_i \notin S$. Since σ_2 is quadratic, these are linear constraints, and they will be independent over \mathbb{Z}_p since the coefficients of $\frac{\partial \sigma_2}{\partial a_i}$ are precisely the entries of the i -th column of the Gram matrix $M_{K_p}^T Q M_{K_p}$, which is nonsingular by our earlier computation. In fact, the same computation reveals that the density of points $\mathbf{a} \in \mathbb{Z}_p^n$ satisfying these conditions will be at most $p^{-(n-m)\lceil \ell \rceil} |\text{Disc}(K_p)|_p^{-1}$. This bound, multiplied by (5.12), is $O(p^{-(n-\frac{m}{2})\ell} |\text{Disc}(K_p)|_p^{-1/2})$.

To obtain the same bound if ℓ is a half-integer (so that $\lfloor \ell \rfloor = \ell - \frac{1}{2}$), we begin by recasting the previous argument slightly. We have seen that, if α meets the conditions described by S , then the integrand in (5.14) is constant in the box $\mathbf{a} + p^{\lfloor \ell \rfloor} \mathbb{Z}_p^{n-m}$, and that the total measure of these boxes is $p^{-(n-m)\lceil \ell \rceil} |\text{Disc}(K_p)|_p^{-1}$.

We now slightly enlarge the boxes, dividing \mathbb{Z}_p^{n-m} into boxes of the form $\mathbf{a} + p^{\lfloor \ell \rfloor} \mathbb{Z}_p^{n-m}$. Here a choice of the coordinates in S will be fixed, and each \mathbf{a} will be chosen to meet the conditions described by S and satisfy $\psi_{p^{2k}}(f_{\mathbf{a}}) = 1$. The following will hold for all $\mathbf{a}' = \mathbf{a} + p^{\lfloor \ell \rfloor} \mathbf{b} \in \mathbf{a} + p^{\lfloor \ell \rfloor} \mathbb{Z}_p^{n-m}$:

1. We will have $|\text{Disc}(f_{\mathbf{a}'})|_p \leq |\text{Disc}(f_{\mathbf{a}})|_p$, with strict inequality if and only if \mathbf{b} satisfies any of a nonempty finite set of \mathbb{Z}_p -linear constraints \pmod{p} .
2. $\psi_{p^{2k}}(f_{\mathbf{a}})$ will be identically 1.
3. The quantity $-u_1 \sigma_1(\boldsymbol{\lambda}) + u_2 \sigma_2(\boldsymbol{\lambda}) \pmod{p^{2k}}$ will be constant on boxes of side length $p^{\lfloor \ell \rfloor}$, and hence may be written as the sum of a constant depending only on \mathbf{a} , and a quadratic polynomial $q(\mathbf{b})$ depending only on $\mathbf{b} \pmod{p}$.
4. After an invertible \mathbb{Z}_p -linear change of variables, we may write $q(\mathbf{b}) = c_1 b_1^2 + \dots + c_t b_t^2 + (\text{terms involving only } b_{t+1}, \dots, b_{n-m})$, where we have $n-m-t \leq \text{corank}_{\mathbb{F}_p}(\sigma_2)$.

Ignoring the condition (1) for now, the integral over $\mathbf{a} + p^{\lfloor \ell \rfloor} \mathbf{b}$ reduces to $|\text{Disc}(f_{\mathbf{a}})|_p^{1/2} p^{-(n-m)\lceil \ell \rceil}$ times a sum over $\mathbf{b} \pmod{p} \in \mathbb{F}_p^{n-m}$, which factors as a product of t Gauss sums, bounded above by $p^{t/2}$, and a sum of the other $n-m-t$ variables, which is $\leq p^{n-m-t} \leq p^{\text{corank}_{\mathbb{F}_p}(\sigma_2)} \ll |\text{Disc}(K_p)|_p^{-1}$. Altogether we obtain a bound

$$\begin{aligned} &\ll |\text{Disc}(f_{\mathbf{a}})|_p^{1/2} p^{-(n-m)\lceil \ell \rceil} p^{t/2} |\text{Disc}(K_p)|_p^{-1} \\ &\ll |\text{Disc}(f_{\mathbf{a}})|_p^{1/2} p^{-(n-m)\ell} |\text{Disc}(K_p)|_p^{-1} \\ &\ll p^{-(n-\frac{m}{2})\ell} |\text{Disc}(K_p)|_p^{-1/2}. \end{aligned} \tag{5.15}$$

We handle the condition (1) by an inductive argument. Passing to the set of \mathbf{b} satisfying any one of the linear constraints, and choosing a \mathbb{Z}_p -basis for this set, (1)-(4) will remain true, with the rank t of the Gauss sum either remaining the same, or decreasing by 1. Since $|\text{Disc}(f)|_p$ decreases by a factor of p , as does the measure of the set being integrated over, while t may decrease by 1, we obtain a bound a factor of $O(p^{-1})$ smaller than (5.15). Continuing to pass to such subsets, and using inclusion-exclusion as needed, we obtain the same bound (5.15) incorporating (1).

Our density bounds for $a_i \in S$ and $a_i \notin S$ have not been proved independent, so adding the minimum of (5.13) and (5.15) over the 2^n choices for S , and noting that $|\text{Disc}(K_p)|_p \leq 1$, we obtain

$$\frac{|\text{Disc}(K_p)|_p^{1/2}}{|\text{Aut}(K_p)|} \int_{\mathcal{O}_{K_p}} |\text{Disc}(f_{\alpha})|_p^{1/2} \psi_{p^{2k}}(f_{\alpha}) e_{p^{2k}}(-u_1 \sigma_1(\boldsymbol{\lambda}) + u_2 \sigma_2(\boldsymbol{\lambda})) d\mu(\alpha) \ll_n p^{-\frac{2n\ell}{3}}.$$

Adding this across the $O_n(1)$ étale algebras, and recalling that $\ell = k - \frac{v_P(u_2)}{2}$, we deduce the result. \square

6 Polynomials with a large and powerful index

The main goal of this section is to prove Theorem 2.6, which states that the number of polynomials for which $\text{index}(f) \gg_n H^{\frac{n(n-3)}{2}}$ but $\text{rad}(\text{index}(f)) < H^{1-\epsilon}$ is $O_{n,\epsilon}(H^{\frac{n^2+n}{2}-\frac{4}{3}-\frac{4}{n}+\epsilon} + H^{\frac{n^2+n}{2}-\frac{2n}{3}+3+\epsilon})$.

We begin with the following elementary lemma that, in the context of Theorem 2.6, will allow us to choose a convenient divisor of the index.

Lemma 6.1. *Let $m \geq 2$ be an integer, and let $C = \text{rad}(m)$ be the product of the primes dividing m . Let $k \geq 2$. If $m \geq C^{2k-2}$, then for every $x \in \mathbb{R}$ such that $C^{k-1} \leq x \leq m/C^{k-1}$, m has a k -powerful divisor d in the interval $[x, Cx]$.*

(Recall that an integer d is k -powerful if every prime $p \mid d$ divides d to order at least k .)

Proof. We begin by proving a slightly stronger statement in the case that m is itself k -powerful, in particular producing a k -powerful divisor in every interval of the form $[x, Cx]$ with $C^{k-1} \leq x \leq m/C$. If $x \leq C^k$, then we simply take the divisor d to be C^k . If $x > C^k$, then we consider divisors of the form $C^k a$ with a a divisor of m/C^k , and we claim there must be such a divisor a in the interval $[x/C^k, x/C^{k-1}]$. If this interval includes m/C^k , then we take the divisor a to be m/C^k itself. Otherwise, let a be the minimal divisor of m/C^k greater than x/C^{k-1} . Then every prime divisor p of a is at most C , and thus $a/p \geq x/C^k$. On the other hand, by the minimality assumption on a , $a/p \leq x/C^{k-1}$, and thus a/p is the claimed divisor. This completes the proof in the case m is k -powerful.

If m is not k -powerful, let m' denote the maximal k -powerful divisor of m and let $C' = \text{rad}(m')$, and notice that $C' \leq C$. Additionally, we have $m/C^{k-1} \leq m'/C'^{(k-1)}$. It then follows from the previous paragraph that m' has a k -powerful divisor in the interval $[x, C'x] \subseteq [x, Cx]$ for x as in the statement of the lemma, and thus m must too. \square

We now turn to the proof of Theorem 2.6.

Proof of Theorem 2.6. We wish to show that the number of monic polynomials $f(x) \in \mathbb{Z}[x]$ of degree n and height H for which $\text{rad}(\text{index}(f)) < H^{1-\epsilon}$ but $\text{index}(f) > H^{\frac{n(n-3)}{2}}$ is $O_{n,\epsilon}(H^{\frac{n^2+n}{2}-\frac{4}{3}-\frac{4}{n}+\epsilon} + H^{\frac{n^2+n}{2}-\frac{2n}{3}+3+\epsilon})$. Suppose we are considering such a polynomial f . By Lemma 6.1 applied with $k = 3$, there is some cubefull divisor d of the index of f satisfying $H^{2-2\epsilon} < d \leq H^{3-3\epsilon}$. By the remainder theorem, given such a divisor d , the polynomial f will lie in the support of the function $\psi_{d^2} = \prod_{p^k \parallel d} \psi_{p^{2k}}$. Moreover, the support of this function is determined by congruence conditions $(\text{mod } d^2)$. Since $d^2 > H$, we may not simply estimate the number of such polynomials by naïvely counting the number in each residue class. Instead, we identify the space of monic, degree n polynomials with \mathbb{Z}^n and we let $\phi: \mathbb{R}^n \rightarrow \mathbb{R}$ be a non-negative Schwartz function, chosen so that it is greater than 1 on the unit box $[-1, 1]^n$ and so that its Fourier transform $\widehat{\phi}$ has compact support contained in $[-\frac{1}{2n^5}, \frac{1}{2n^5}]^n$.

Then by Poisson summation,

$$\begin{aligned} \sum_{\mathbf{c} \in \mathbb{Z}^n} \phi\left(\frac{c_1}{H}, \dots, \frac{c_n}{H^n}\right) \psi_{d^2}(f_{\mathbf{c}}) &= H^{\frac{n^2+n}{2}} \sum_{\mathbf{u} \in \mathbb{Z}^n} \widehat{\phi}\left(\frac{u_1 H}{d^2}, \dots, \frac{u_n H^n}{d^2}\right) \widehat{\psi}_{d^2}(\mathbf{u}) \\ &= H^{\frac{n^2+n}{2}} \widehat{\psi}_{d^2}(\mathbf{0}) \widehat{\phi}(\mathbf{0}) + H^{\frac{n^2+n}{2}} \sum_{\mathbf{0} \neq \mathbf{u} \in \mathbb{Z}^n} \widehat{\phi}\left(\frac{u_1 H}{d^2}, \dots, \frac{u_n H^n}{d^2}\right) \widehat{\psi}_{d^2}(\mathbf{u}) \end{aligned} \quad (6.1)$$

where $\widehat{\psi}_{d^2}(\mathbf{u})$ is defined by

$$\widehat{\psi}_{d^2}(\mathbf{u}) = \frac{1}{d^{2n}} \sum_{f \in (\mathbb{Z}/d^2\mathbb{Z})^n} \psi_d(f) e^{\frac{2\pi i \langle f, \mathbf{u} \rangle}{d^2}}.$$

By the remainder theorem, $\widehat{\psi}_{d^2}(\mathbf{u})$ may be decomposed as $\widehat{\psi}_{d^2}(\mathbf{u}) = \prod_{p^k \parallel d} \widehat{\psi}_{p^{2k}}(\gamma_p \mathbf{u})$ for some γ_p coprime to p . Thus, $\widehat{\psi}_{d^2}(\mathbf{u})$ may be bounded by the results of the previous section, in particular Lemma 5.1, Lemma 5.4, and Lemma 5.6. Using Lemma 5.1, the first term on the right-hand side of (6.1) is $O_{n, \phi, \epsilon}(H^{\frac{n^2+n}{2}} d^{-1-\frac{2}{n}+\epsilon})$. Added over cubefull integers $d > H^{2-2\epsilon}$, this yields a total contribution that is $O_{n, \phi, \epsilon}(H^{\frac{n^2+n}{2}-\frac{4}{3}-\frac{4}{n}+\epsilon})$, which matches the bound claimed in Theorem 2.6.

For the contribution from the sum of the non-trivial Fourier coefficients, we first note that by the compact support of $\widehat{\phi}$, the summation is supported on those \mathbf{u} for which each $|u_i| \leq d^2/2H^i$. Since we have assumed that $d \leq H^{3-3\epsilon}$, this implies that $u_i = 0$ for $i \geq 6$. In particular, since $u_6 = 0$, it follows from the arithmetic progression property of the support (Lemma 5.4) of $\widehat{\psi}_{d^2}$ that each u_i for $i \leq 5$ must be divisible by $d/\gcd(n^{6-i}, d)$. (In the worst case, when (5.6) holds for every prime p dividing d , we bound $b \leq v_p(n)$). Thus, writing $u_i = (d/\gcd(n^{6-i}, d))u'_i$, we find that u'_i satisfies $|u'_i| \leq d/H^i$, which in particular also implies that $u_i = 0$ for $i \geq 3$. If $u_2 = 0$, then the sum is only over integers $|u_1| \leq d^2/H$. However, by Lemma 5.5, $\widehat{\psi}_{d^2}(\mathbf{u})$ will vanish unless u_1 is divisible by $d^2/\gcd(d^2, n)$, which forces $u_1 = 0$ in this range as well. As we have already considered the contribution from the trivial Fourier coefficient, we may therefore assume that $u_2 \neq 0$.

If $u_2 \neq 0$, then by Lemma 5.6, $\widehat{\psi}_{d^2}(\mathbf{u}) \ll_{n, \epsilon} d^{-2n/3+\epsilon} \gcd(d^2, u_2)^{n/3} \ll_n d^{-n/3+\epsilon} \gcd(d, u'_2)^{n/3}$. This case thus yields a contribution that is

$$\ll_{n, \epsilon} \frac{H^{\frac{n^2+n}{2}+\epsilon}}{d^{n/3}} \sum_{|u'_1| \ll_n d/H} \sum_{0 \neq |u'_2| \ll_n d/H^2} \gcd(d, u'_2)^{n/3} \ll_{n, \epsilon} d H^{\frac{n^2+n}{2}-\frac{2n}{3}-1+\epsilon}, \quad (6.2)$$

by noting that the summation over u'_2 is dominated by the $O(d^\epsilon)$ values u'_2 for which $\gcd(d, u'_2)$ is largest possible, but that this gcd is $\tilde{O}_n(d/H^2)$ from the bound on u'_2 . After summing over cubefull integers $H^{2-2\epsilon} < d < H^{3-3\epsilon}$, we find a total contribution from the non-trivial Fourier coefficients that is $O_{n, \epsilon}(H^{\frac{n^2+n}{2}-\frac{2n}{3}+3+\epsilon})$, which completes the proof of the theorem. \square

7 Strong and weak multiples

In this section, we prove Theorem 2.7, which bounds the number of polynomials of height H whose discriminants have large squarefree divisors. To do this, we build on the work of Bhargava, Shankar, and

Wang [BSW22b]. As before, to any $\mathbf{c} = (c_1, \dots, c_n) \in \mathbb{Z}^n$ we associate the monic degree n polynomial $f_{\mathbf{c}}(x) = x^n + c_1 x^{n-1} + \dots + c_n$.

Theorem 2.7 should be compared with [BSW22b, Theorem 4.4], where a similar statement is proven but where the second term is $H^{\frac{n^2+n}{2} - \frac{1}{5} + \epsilon}$ instead of $H^{\frac{n^2+n}{2} - \frac{1}{2} + \epsilon}$. It should also be compared with [BSW22a, Theorem 4], where a strictly stronger statement is proven. Additionally, as sketched in [BSW22b], it is possible to prove that the first term on the right-hand side is $H^{\frac{n^2+n}{2}}/M$, but we do not pursue this as it is not necessary in the proof of our main theorem.

To prove Theorem 2.7, we follow the same strategy as [BSW22b], offering improvements to each of the two key steps. For a squarefree integer m , let \mathcal{W}_m be the underlying set of polynomials whose discriminant is divisible by m^2 , that is

$$\mathcal{W}_m := \{\mathbf{c} \in \mathbb{Z}^n : |c_i| \leq H^i, m^2 \mid \text{disc}(f_{\mathbf{c}})\}.$$

We decompose this set as follows. For a prime p , we say that $\text{disc}(f)$ is a *strong multiple* of p^2 if $\text{disc}(g)$ is a multiple of p^2 for every polynomial g congruent to $f \pmod{p}$, and we say that $\text{disc}(f)$ is a *weak multiple* of p^2 otherwise. Let $\mathcal{W}_m^{(1)} \subseteq \mathcal{W}_m$ be the subset consisting of polynomials for which $\text{disc}(f)$ is a strong multiple of p^2 for every prime $p \mid m$, and let $\mathcal{W}_m^{(2)} \subseteq \mathcal{W}_m$ be the subset of those for which $\text{disc}(f)$ is a weak multiple of p^2 for every prime $p \mid m$. By considering the factorization of m , Bhargava, Shankar, and Wang [BSW22b] make the simple observation that for any $M > 1$,

$$\bigcup_{\substack{m \geq M \\ m \text{ squarefree}}} \mathcal{W}_m \subseteq \bigcup_{\substack{m \geq \sqrt{M} \\ m \text{ squarefree}}} \mathcal{W}_m^{(1)} \cup \bigcup_{\substack{m \geq \sqrt{M} \\ m \text{ squarefree}}} \mathcal{W}_m^{(2)}.$$

Thus, to prove Theorem 2.7, it is sufficient to prove the following pair of propositions that refine [BSW22b, Theorem 1.5(a)] and [BSW22b, Theorem 1.5(b)], respectively.

Proposition 7.1. *For any $Y > 1$, $H > 0$, and $\epsilon > 0$,*

$$\bigcup_{\substack{m \geq Y \\ m \text{ squarefree}}} \mathcal{W}_m^{(1)} \ll_{n, \epsilon} \frac{H^{\frac{n^2+n}{2} + \epsilon}}{Y} + H^{\frac{n^2+n}{2} - n + 1 + \epsilon}.$$

Proposition 7.2. *For any $Y > 1$ and any $\epsilon > 0$, we have*

$$\bigcup_{\substack{m \geq Y \\ m \text{ squarefree}}} \mathcal{W}_m^{(2)} \ll_n \frac{H^{\frac{n^2+n}{2}}}{Y} + H^{\frac{n^2+n}{2} - \frac{1}{2} + \epsilon}.$$

We now turn to the proofs of these propositions.

7.1 Strong multiples and Proposition 7.1

To prove Proposition 7.1, we apply a line of reasoning similar to the geometric sieve (c.f. Theorem 3.3 of [Bha14]).

Recall that each $\mathbf{c} \in \mathbb{Z}^n$ corresponds to a degree n polynomial $f_{\mathbf{c}} = x^n + c_1 x^{n-1} + \dots + c_n$. The discriminant $\text{disc}(f_{\mathbf{c}})$ is a polynomial in the coefficients c_i . More generally, to any $\mathbf{c} \in \mathbb{R}^n$ we can associate the gradient vector $\mathbf{D}_{\mathbf{c}} := (\frac{\partial \text{disc}(f)}{\partial c_1}, \dots, \frac{\partial \text{disc}(f)}{\partial c_n})$. Then we see that $\text{disc}(f_{\mathbf{c}})$ is a strong multiple of p^2 if and only if both $\text{disc}(f_{\mathbf{c}}) \equiv 0 \pmod{p}$ and $\mathbf{D}_{\mathbf{c}} \equiv \mathbf{0} \pmod{p}$. Thus it suffices to bound the number of \mathbf{c} with $|c_i| \leq H^i$ ($1 \leq i \leq n$) such that $\text{disc}(f_{\mathbf{c}}) \equiv 0 \pmod{q}$ and $\frac{\partial \text{disc}(f)}{\partial c_n} \equiv 0 \pmod{q}$ for some squarefree integer $q \geq Y$. To simplify, we replace $\frac{\partial \text{disc}(f)}{\partial c_n}(\mathbf{c})$ by the resultant of $\frac{\partial \text{disc}(f)}{\partial c_n}(\mathbf{c})$ and $\text{disc}(f_{\mathbf{c}})$ with respect to c_{n-1} . Denote the two polynomials $\text{disc}(f_{\mathbf{c}})$ and $\text{Res}(\text{disc}(f_{\mathbf{c}}), \frac{\partial \text{disc}(f)}{\partial c_n})$ by $g_1(\mathbf{c})$ and $g_2(\mathbf{c})$, respectively, where we notate $g_2(\mathbf{c})$ even though g_2 doesn't depend on c_{n-1} so that it suffices to count common roots \mathbf{c} of g_1 and g_2 .

We first show that $g_1(\mathbf{c})$, as a polynomial in $(\mathbb{Z}[c_1, \dots, c_{n-2}, c_n])[c_{n-1}]$, has degree n and nonzero constant leading coefficient, or equivalently that $g_1(\mathbf{c}) = \sum_{i \leq n} \alpha_i c_{n-1}^i$ with $\alpha_n \neq 0$. To see this, we compute $g_1(\mathbf{c}) = \text{disc}(f_{\mathbf{c}})$ via the resultant matrix A of the monic polynomial $f_{\mathbf{c}}(x)$ and its derivative $f'_{\mathbf{c}}(x)$. In this matrix, we observe that c_{n-1} occurs exactly 0 times in the first $n-1$ rows, twice in rows n to $(2n-2)$, and once in the $(2n-1)$ st row. Thus the degree is at most n . We explicitly compute that the coefficient α_n is

$$\alpha_n = \sum_{0 \leq k \leq n-1} (-n)^k \cdot \binom{n-1}{k} = (1-n)^{n-1} \neq 0.$$

Next, from the weighted homogeneity property of resultants, we observe that $g_2(\mathbf{c})$ may be expressed as a homogeneous polynomial of degree $n^2(n-2)$ in the roots of f , and is therefore a weighted homogeneous polynomial in \mathbf{c} of the same degree, where each c_i has weight i . It therefore has degree at most $n(n-2)$ as a polynomial in c_n , and we observe that there is a unique permutation giving rise to a term of such degree in the determinant representation of the resultant, corresponding to choosing the top entry of the columns associated to $\text{disc}(f_{\mathbf{c}})$ and the bottom entry of the columns associated to $\frac{\partial \text{disc}(f)}{\partial c_n}$. We conclude that $g_2(\mathbf{c})$, as a polynomial in c_n , has degree $n(n-2)$ and that its leading coefficient is a non-zero constant depending only on n . In particular, $g_2(\mathbf{c})$ will be a non-zero polynomial in c_n for any choice of the coordinates (c_1, \dots, c_{n-2}) .

The total number of tuples among the first $n-2$ coordinates (c_1, \dots, c_{n-2}) satisfying $|c_i| \leq H^i$ is of the order $H^{n(n+1)/2 - (2n-1)}$. For each such tuple, the number of c_n such that $g_2(c_1, \dots, c_{n-2}, c_n) = 0$ is absolutely bounded by the degree in c_n , and thus bounded by $O_n(1)$; the number of c_{n-1} is trivially bounded by H^{n-1} . These contribute at most $O_n(H^{n(n+1)/2 - n})$ many \mathbf{c} .

If $g_2(c_1, \dots, c_{n-2}, c_n) \neq 0$, then we take $q \geq Y$ to be any squarefree divisor of $g_2(c_1, \dots, c_{n-2}, c_n)$; there are at most $O_{n,\epsilon}(H^\epsilon)$ many such choices q . The number of c_{n-1} with $|c_{n-1}| \leq H^{n-1}$ such that $g_1(c_1, \dots, c_n) \equiv 0 \pmod{q}$ is at most $O_n(\max\{1, H^{n-1}/Y\})$. In total, these contribute at most $O_n(H^{n(n+1)/2 + \epsilon}/Y) + O_n(H^{n(n+1)/2 + \epsilon - (n-1)})$.

Combining these two bounds completes the proof of Proposition 7.1. \square

7.2 Weak multiples and Proposition 7.2

We now turn to Proposition 7.2, which is an improvement of the second error term in [BSW22b, Theorem 1.5(b)]. Our input consists simply of the application of a sharper sieve result at the heart of their argument. In particular, the error term in question originates in Propositions 2.6 and 3.5 of [BSW22b], depending on whether n is odd or even. The proofs of both propositions rely on an application of the Selberg sieve

to bound the number of “distinguished orbits” of an appropriate orthogonal group acting on the space of symmetric $n \times n$ matrices over $\frac{1}{2}\mathbb{Z}$. For each prime p , they establish that they are able to sieve out a proportion of congruence conditions which is uniformly bounded away from 0, and then an application of the Selberg sieve yields their result.

It turns out that an application of the large sieve in the form of [Ser08, Theorem 10.1.1] immediately yields the stated stronger result. We also note that further improvements might be possible, in particular by incorporating Fourier analysis when applying the Selberg sieve as a large sieve, as done in [AGLO⁺23].

Acknowledgements

This collaboration arose initially from an AIM workshop on Fourier Analysis, Arithmetic Statistics, and Discrete Restriction organized by two of the authors (T.A. and F.T.) and Trevor Wooley. The authors thank AIM for a very supportive working environment. The authors thank Manjul Bhargava, Arul Shankar, and Jerry Wang for sharing a preprint version of their manuscript [BSW22a], and for many useful conversations. They would also like to thank the anonymous referee for their many useful comments.

References

- [AGLO⁺23] Theresa C. Anderson, Ayla Gafni, Robert J. Lemke Oliver, David Lowry-Duda, George Shakan, and Ruixiang Zhang. Quantitative Hilbert irreducibility and almost prime values of polynomial discriminants. *Int. Math. Res. Not. IMRN*, (3):2188–2214, 2023. [22](#)
- [Bha05] Manjul Bhargava. The density of discriminants of quartic rings and fields. *Ann. of Math.* (2), 162(2):1031–1063, 2005. [2](#)
- [Bha10] Manjul Bhargava. The density of discriminants of quintic rings and fields. *Ann. of Math.* (2), 172(3):1559–1591, 2010. [2](#)
- [Bha14] Manjul Bhargava. The geometric sieve and the density of squarefree values of invariant polynomials, 2014. arXiv:1402.0031. [20](#)
- [BSW22a] Manjul Bhargava, Arul Shankar, and Xiaoheng Wang. An improvement on Schmidt’s bound on the number of number fields of bounded discriminant and small degree. *Forum Math. Sigma*, 10:Paper No. e86, 13, 2022. [2](#), [20](#), [22](#)
- [BSW22b] Manjul Bhargava, Arul Shankar, and Xiaoheng Wang. Squarefree values of polynomial discriminants I. *Invent. Math.*, 228(3):1037–1073, 2022. [2](#), [3](#), [20](#), [21](#)
- [CDyDO02] Henri Cohen, Francisco Diaz y Diaz, and Michel Olivier. Enumerating quartic dihedral extensions of \mathbb{Q} . *Compositio Math.*, 133(1):65–93, 2002. [2](#)
- [Cou20] Jean-Marc Couveignes. Enumerating number fields. *Ann. of Math.* (2), 192(2):487–497, 2020. [2](#)
- [Dav51] H. Davenport. On a principle of Lipschitz. *J. London Math. Soc.*, 26:179–183, 1951. [9](#)

- [DH69] H. Davenport and H. Heilbronn. On the density of discriminants of cubic fields. *Bull. London Math. Soc.*, 1:345–348, 1969. [2](#)
- [EV06] Jordan S. Ellenberg and Akshay Venkatesh. The number of extensions of a number field with fixed degree and bounded discriminant. *Ann. of Math. (2)*, 163(2):723–741, 2006. [2](#)
- [GKZ08] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky. *Discriminants, resultants and multidimensional determinants*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2008. Reprint of the 1994 edition. [11](#)
- [LOT22] Robert J. Lemke Oliver and Frank Thorne. Upper bounds on number fields of given degree and bounded discriminant. *Duke Math. J.*, 171(15):3077–3087, 2022. [2](#)
- [Sch95] Wolfgang M. Schmidt. Number fields of given degree and bounded discriminant. *Astérisque*, 4(228):189–195, 1995. Columbia University Number Theory Seminar (New York, 1992). [1](#), [3](#)
- [Ser08] Jean-Pierre Serre. *Topics in Galois theory*, volume 1 of *Research Notes in Mathematics*. A K Peters, Ltd., Wellesley, MA, second edition, 2008. With notes by Henri Darmon. [7](#), [22](#)
- [ST23] Arul Shankar and Jacob Tsimerman. Heuristics for the asymptotics of the number of S_n -number fields. *J. Lond. Math. Soc. (2)*, 107(5):1613–1634, 2023. [2](#), [7](#)

AUTHORS

Theresa C. Anderson
 Gregg Zeitlin Assistant Professor of Mathematical Sciences
 Carnegie Mellon University
 Pittsburgh, PA
 tanders2@andrew.cmu.edu
<https://faculty.math.cmu.edu/anderson/>

Ayla Gafni
 Assistant Professor
 University of Mississippi
 Oxford, MS
 argafni@olemiss.edu
<https://math.olemiss.edu/ayla-gafni/>

Kevin Hughes
Lecturer
Edinburgh Napier University
Edinburgh, Scotland
khughes.math@gmail.com
<https://sites.google.com/site/khughesmath/>

Robert J. Lemke Oliver
Associate Professor
Tufts University
Medford, MA
robert.lemke_oliver@tufts.edu
<https://rlemke01.math.tufts.edu/>

David Lowry-Duda
Senior Research Scientist
ICERM
Providence, RI
david@lowryduda.com
<https://davidlowryduda.com/>

Frank Thorne
Professor
University of South Carolina
Columbia, SC
thorne@math.sc.edu
<https://thornef.github.io/>

Jiuya Wang
Assistant Professor
University of Georgia
Athens, GA
jiuya.wang@uga.edu
<https://wangjiuya.github.io/>

Ruixiang Zhang
Assistant Professor
University of California, Berkeley
Berkeley, CA
ruixiang@berkeley.edu
<https://sites.google.com/view/ruixiang-zhang/home>